SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Davis, Roger J.
Raingeaud, Joel
Gupta, Shashi
Derijard, Benoit

(ii) TITLE OF INVENTION: CYTOKINE-, STRESS-, AND ONCOPROTEIN-ACTIVATED HUMAN PROTEIN KINASE KINASES

(iii) NUMBER OF SEQUENCES: 16

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 225 Franklin Street
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02110

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WordPerfect 5.1

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/446,083
(B) FILING DATE: 05/19/95

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Fasse, J. Peter
(B) REGISTRATION NUMBER: 32,983
(C) REFERENCE/DOCKET NUMBER: 04020/09001

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 542-5070
(B) TELEFAX: (617) 542-8906

(2) INFORMATION FOR SEQ ID NO:1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2030 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
TGGCTGGCAA TGGCCTTGCT GACCTCGAGC CGGGCCCACG TGGGGACCTT TGGAGCACAG   60
CCTACGATCC TGGTGCAAGG CCGGTGGATG CAGAGGCCAG TCCATATACC ACCCAGGCCT  120
GCGAGGAGCG TGGTCCCCAC CCATCCAGCC CATATGTGCA AGTGCCCTTG ACAGAGAGGC  180
TGGTCATATC CATGGTGACC ATTTATGGGC CACAACAGGT CCCCATCTGC GCAGTGAACC  240
CTGTGCTGAG CACCTTGCAG ACGTGATCTT GCTTCGTCCT GCAGCACTGT GCGGGGCAGG  300
AAAATCCAAG AGGAAGAAGG ATCTACGGAT ATCCTGCATG TCCAAGCCAC CCGCACCCAA  360
CCCCACACCC CCCCGGAACC TGGACTCCCG GACCTTCATC ACCATTGGAG ACAGAAACTT  420
TGAGGTGGAG GCTGATGACT TGGTGACCAT CTCAGAACTG GGCCGTGGAG CCTATGGGGT  480
GGTAGAGAAG GTGCGGCACG CCCAGAGCGG CACCATCATG GCCGTGAAGC GGATCCGGGC  540
CACCGTGAAC TCACAGGAGC AGAAGCGGCT GCTCATGGAC CTGGACATCA ACATGCGCAC  600
GGTCGACTGT TTCTACACTG TCACCTTCTA CGGGGCACTA TTCAGAGAGG GAGACGTGTG  660
GATCTGCATG GAGCTCATGG ACACATCCTT GGACAAGTTC TACCGGAAGG TGCTGGATAA  720
AAACATGACA ATTCCAGAGG ACATCCTTGG GGAGATTGCT GTGTCTATCG TGCGGGCCCT  780
GGAGCATCTG CACAGCAAGC TGTCGGTGAT CCACAGAGAT GTGAAGCCCT CCAATGTCCT  840
TATCAACAAG GAGGGCCATG TGAAGATGTG TGACTTTGGC ATCAGTGGCT ACTTGGTGGA  900
CTCTGTGGCC AAGACGATGG ATGCCGGCTG CAAGCCCTAC ATGGCCCCTG AGAGGATCAA  960
CCCAGAGCTG AACCAGAAGG GCTACAATGT CAAGTCCGAC GTCTGGAGCC TGGGCATCAC 1020
CATGATTGAG ATGGCCATCC TGCGGTTCCC TTACGAGTCC TGGGGGACCC CGTTCCAGCA 1080
GCTGAAGCAG GTGGTGGAGG AGCCGTCCCC CCAGCTCCCA GCCGACCGTT TCTCCCCCGA 1140
GTTTGTGGAC TTCACTGCTC AGTGCCTGAG GAAGAACCCC GCAGAGCGTA TGAGCTACCT 1200
GGAGCTGATG GAGCACCCCT TCTTCACCTT GCACAAAACC AAGAAGACGG ACATTGCTGC 1260
CTTCGTGAAG AAGATCCTGG GAGAAGACTC ATAGGGGCTG GGCCTCGGAC CCCACTCCGG 1320
CCCTCCAGAG CCCCACAGCC CCATCTGCGG GGGCAGTGCT CACCCACACC ATAAGCTACT 1380
GCCATCCTGG CCCAGGGCAT CTGGGAGGAA CCGAGGGGGC TGCTCCCACC TGGCTCTGTG 1440
GCGAGCCATT TGTCCCAAGT GCCAAAGAAG CAGACCATTG GGGCTCCCAG CCAGGCCCTT 1500
GTCGGCCCCA CCAGTGCCTC TCCCTGCTGC TCCTAGGACC CGTCTCCAGC TGCTGAGATC 1560
CTGGACTGAG GGGGCCTGGA TGCCCCCTGT GGATGCTGCT GCCCCTGCAC AGCAGGCTGC 1620
CAGTGCCTGG GTGGATGGGC CACCGCCTTG CCCAGCCTGG ATGCCATCCA AGTTGTATAT 1680
TTTTTTAATC TCTCGACTGA ATGGACTTTG CACACTTTGG CCCAGGGTGG CCACACCTCT 1740
ATCCCGGCTT TGGTGCGGGG TACACAAGAG GGGATGAGTT GTGTGAATAC CCCAAGACTC 1800
CCATGAGGGA GATGCCATGA GCCGCCCAAG GCCTTCCCCT GGCACTGGCA AACAGGGCCT 1860
CTGCGGAGCA CACTGGCTCA CCCAGTCCTG CCCGCCACCG TTATCGGTGT CATTCACCTT 1920
TCGTGTTTTT TTTAATTTAT CCTCTGTTGA TTTTTTCTTT TGCTTTATGG GTTTGGCTTG 1980
TTTTTCTTGC ATGGTTTGGA GCTGATCGCT TCTCCCCCAC CCCCTAGGGG             2030
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 363 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
                5                  10                  15
Met Ser Lys Pro Pro Ala Pro Asn Pro Thr Pro Pro Arg Asn Leu Asp
            20                  25                  30
Ser Arg Thr Phe Ile Thr Ile Gly Asp Arg Met Phe Glu Val Glu Ala
        35                  40                  45
Asp Asp Leu Val Thr Ile Ser Glu Leu Gly Arg Gly Ala Tyr Gly Val
    50                  55                  60
```

Val Glu Lys Val Arg His Ala Gln Ser Gly Thr Ile Met Ala Val Lys
65 70 75 80
Arg Ile Arg Ala Thr Val Asn Ser Gln Glu Gln Lys Arg Leu Leu Met
85 90 95
Asp Leu Asp ILe Asn Met Arg Thr Val Asp Cys Phe Tyr Thr Val Thr
100 105 110
Phe Tyr Gly Ala Leu Phe Arg Glu Gly Asp Val Trp Ile Cys Met Glu
115 120 125
Leu Met Asp Thr Ser Leu Asp Lys Phe Tyr Arg Lys Val Leu Asp Lys
130 135 140
Asn Met Thr Ile Pro Glu Asp Ile Leu Gly Glu Ile Ala Val Ser Ile
145 150 155 160
Val Arg Ala Leu Glu His Leu His Ser Lys Leu Ser Val Ile His Arg
165 170 175
Asp Val Lys Pro Ser Asn Val Leu Ile Asn Lys Glu Gly His Val Lys
200 205 210
Met Cys Asp Phe Gly Ile Ser Gly Tyr Leu Val Asp Ser Val Ala Lys
215 220 225
Thr Met Asp Ala Gly Cys Lys Pro Tyr Met Ala Pro Glu Arg Ile Asn
230 235 240
Pro Glu Leu Asn Gln Lys Gly Tyr Asn Val Lys Ser Asp Val Trp Ser
245 250 255 260
Leu Gly Ile Thr Met Ile Glu Met Ala Ile Leu Arg Phe Pro Tyr Glu
265 270 275
Ser Trp Gly Thr Pro Phe Gln Gln Leu Lys Gln Val Val Glu Glu Pro
280 285 290
Ser Pro Gln Leu Pro Ala Asp Arg Phe Ser Pro Glu Phe Val Asp Phe
295 300 305
Thr Ala Gln Cys Leu Arg Lys Asn Pro Ala Glu Arg Met Ser Tyr Leu
310 315 320
Glu Leu Met Glu His Pro Phe Phe Thr Leu His Lys Thr Lys Lys Thr
325 330 335 338
Asp Ile Ala Ala Phe Val Lys Lys Ile Leu Gly Glu Asp Ser

(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: DNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TAGCTGCAGC ACAGCCTTCC CTAACGTTGC AACTGGGGGA AAAATCACTT TCCAGTCTGT 60
TTTGCAAGGT GTGCATTTCC ATCTTGATTC CCTGAAAGTC CATCTGCTGC ATCGGTCAAG 120
AGAAACTCCA CTTGCATGAA GATTGCACGC CTGCAGCTTG CATCTTTGTT GCAAAACTAG 180
CTACAGAAGA GAAGCAAGGC AAAGTCTTTT GTGCTCCCCT CCCCCATCAA AGGAAAGGGG 240
AAAATGTCTC AGTCGAAAGG CAAGAAGCGA AACCCTGGCC TTAAAATTCC AAAAGAAGCA 300
TTTGAACAAC CTCAGACCAG TTCCACACCA CCTAGAGATT TAGACTCCAA GGCTTGCATT 360
TCTATTGGAA ATCAGAACTT TGAGGTGAAG GCAGATGACC TGGAGCCTAT AATGGAACTG 420

```
GGACGAGGTG CGTACGGGGT GGTGGAGAAG ATGCGGCACG TGCCCAGCGG GCAGATCATG  480
GCAGTGAAGC GGATCCGAGC CACAGTAAAT AGCCAGGAAC AGAAACGGCT ACTGATGGAT  540
TTGGATATTT CCATGAGGAC GGTGGACTGT CCATTCACTG TCACCTTTTA TGGCGCACTG  600
TTTCGGGAGG GTGATGTGTG GATCTGCATG GAGCTCATGG ATACATCACT AGATAAATTC  660
TACAAACAAG TTATTGATAA AGGCCAGACA ATTCCAGAGG ACATCTTAGG GAAAATAGCA  720
GTTTCTATTG TAAAAGCATT AGAACATTTA CATAGTAAGC TGTCTGTCAT TCACAGAGAC  780
GTCAAGCCTT CTAATGTACT CATCAATGCT CTCGGTCAAG TGAAGATGTG CGATTTTGGA  840
ATCAGTGGCT ACTTGGTGGA CTCTGTTGCT AAAACAATTG ATGCAGGTTG CAAACCATAC  900
ATGGCCCCTG AAAGAATAAA CCCAGAGCTC AACCAGAAGG GATACAGTGT GAAGTCTGAC  960
ATTTGGAGTC TGGGCATCAC GATGATTGAG TTGGCCATCC TTCGATTTCC CTATGATTCA 1020
TGGGGAACTC CATTTCAGCA GCTCAAACAG GTGGTAGAGG AGCCATCGCC ACAACTCCCA 1080
GCAGACAAGT TCTCTGCAGA GTTTGTTGAC TTTACCTCAC AGTGCTTAAA GAAGAATTCC 1140
AAAGAACGGC CTACATACCC AGAGCTAATG CAACATCCAT TTTTCACCCT ACATGAATCC 1200
AAAGGAACAG ATGTGGCATC TTTTGTAAAA CTGATTCTTG GAGACTAAAA AGCAGTGGAC 1260
TTAATCGGTT GACCCTACTG TGGATTGGTG GGTTTCGGGG TGAAGCAAGT TCACTACAGC 1320
ATCAATAGAA AGTCATCTTT GAGATAATTT AACCCTGCCT CTCAGAGGGT TTTCTCTCCC 1380
AATTTTCTTT TTACTCCCCC TCTTAAGGGG GCCTTGGAAT CTATAGTATA GAATGAACTG 1440
TCTAGATGGA TGAATTATGA TAAAGGCTTA GGACTTCAAA AGGTGATTAA ATATTTAATG 1500
ATGTGTCATA TGAGTCCTCA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1560
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA                    1602
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 344 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
                  5                  10                  15
Met Ser Gln Ser Lys Gly Lys Lys Arg Asn Pro Gly Leu Lys Ile Pro
             20                  25                  30
Lys Glu Ala Phe Glu Gln Pro Gln Thr Ser Ser Thr Pro Pro Arg Asp
         35                  40                  45
Leu Asp Ser Lys Ala Cys Ile Ser Ile Gly Asn Gln Asn Phe Glu Val
     50                  55                  60
Lys Ala Asp Asp Leu Glu Pro Ile Met Glu Leu Gly Arg Gly Ala Tyr
 65                  70                  75                  80
Gly Val Val Glu Lys Met Arg His Val Pro Ser Gly Gln Ile Met Ala
                 85                  90                  95
Val Lys Arg Ile Arg Ala Thr Val Asn Ser Gln Glu Gln Lys Arg Leu
            100                 105                 110
Leu Met Asp Leu Asp Ile Ser Met Arg Thr Val Asp Cys Pro Phe Thr
        115                 120                 125
Val Thr Phe Tyr Gly Ala Leu Phe Arg Glu Gly Asp Val Trp Ile Cys
    130                 135                 140
Met Glu Leu Met Asp Thr Ser Leu Asp Lys Phe Tyr Lys Gln Val Ile
145                 150                 155                 160
Asp Lys Gly Gln Thr Ile Pro Glu Asp Ile Leu Gly Lys Ile Ala Val
```

```
                 165                 170                 175
Ser Ile Val Lys Ala Leu Glu His Leu His Ser Lys Leu Ser Val Ile
            180                 185                 190
His Arg Asp Val Lys Pro Ser Asn Val Leu Ile Asn Ala Leu Gly Gln
        195                 200                 205
Val Lys Met Cys Asp Phe Gly Ile Ser Gly Tyr Leu Val Asp Ser Val
    210                 215                 220
Ala Lys Thr Ile Asp Ala Gly Cys Lys Pro Tyr Met Ala Pro Glu Arg
225                 230                 235                 240
Ile Asn Pro Glu Leu Asn Gln Lys Gly Tyr Ser Val Lys Ser Asp Ile
                245                 250                 255
Trp Ser Leu Gly Ile Thr Met Ile Glu Leu Ala Ile Leu Arg Phe Pro
            260                 265                 270
Tyr Asp Ser Trp Gly Thr Pro Phe Gln Gln Leu Lys Gln Val Val Glu
        275                 280                 285
Glu Pro Ser Pro Gln Leu Pro Ala Asp Lys Phe Ser Ala Glu Phe Val
    290                 295                 300
Asp Phe Thr Ser Gln Cys Leu Lys Lys Asn Ser Lys Glu Arg Pro Thr
305                 310                 315                 320
Tyr Pro Glu Leu Met Gln His Pro Phe Phe Thr Leu His Glu Ser Lys
                325                 330                 334
Gly Thr Asp Val Ala Ser Phe Val Lys Leu Ile Leu Gly Asp
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3498 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CTAGGGTCCC CGGCGCCAGG CCACCCGGCC GTCAGCAGCA TGCAGGGTAA ACGCAAAGCA   60
CTGAAGTTGA ATTTTGCAAA TCCACCTTTC AAATCTACAG CAAGGTTTAC TCTGAATCCC  120
AATCCTACAG GAGTTCAAAA CCCACACATA GAGAGACTGA GAACACACAG CATTGAGTCA  180
TCAGGAAAAC TGAAGATCTC CCCTGAACAA CACTGGGATT TCACTGCAGA GGACTTGAAA  240
GACCTTGGAG AAATTGGACG AGGAGCTTAT GGTTCTGTCA ACAAAATGGT CCACAAACCA  300
AGTGGGCAAA TAATGGCAGT TAAAAGAATT CGGTCAACAG TGGATGAAAA AGAACAAAAA  360
CAACTTCTTA TGGATTTGGA TGTAGTAATG CGGAGTAGTG ATTGCCCATA CATTGTTCAG  420
TTTTATGGTG CACTCTTCAG AGAGGGTGAC TGTTGGATCT GTATGGAACT CATGTCTACC  480
TCGTTTGATA AGTTTTACAA ATATGTATAT AGTGTATTAG ATGATGTTAT TCCAGAAGAA  540
ATTTTAGGCA AAATCACTTT AGCAACTGTG AAAGCACTAA ACCACTTAAA AGAAAACTTG  600
AAAATTATTC ACAGAGATAT CAAACCTTCC AATATTCTTC TGGACAGAAG TGGAAATATT  660
AAGCTCTGTG ACTTCGGCAT CAGTGGACAG CTTGTGGACT CTATTGCCAA GACAAGAGAT  720
GCTGGCTGTA GGCCATACAT GGCACCTGAA AGAATAGACC CAAGCGCATC ACGACAAGGA  780
TATGATGTCC GCTCTGATGT CTGGAGTTTG GGGATCACAT TGTATGAGTT GGCCACAGGC  840
CGATTTCCTT ATCCAAAGTG GAATAGTGTA TTTGATCAAC TAACACAAGT CGTGAAAGGA  900
GATCCTCCGC AGCTGAGTAA TTCTGAGGAA AGGGAATTCT CCCCGAGTTT CATCAACTTT  960
GTCAACTTGT GCCTTACGAA GGATGAATCC AAAAGGCCAA AGTATAAAGA GCTTCTGAAA 1020
CATCCCTTTA TTTTGATGTA TGAAGAACGT GCCGTTGAGG TCGCATGCTA TGTTTGTAAA 1080
```

ATCCTGGATC AAATGCCAGC TACTCCCAGC TCTCCCATGT ATGTCGATTG ATATCGYTGC 1140
TACATCAGAC TCTAGAAAAA AGGGCTGAGA GGAAGCAAGA CGTAAAGAAT TTTCATCCCG 1200
TATCACAGTG TTTTTATTGC TCGCCCAGAC ACCATGTGCA ATAAGATTGG TGTTCGTTTC 1260
CATCATGTCT GTATACTCCT GTCACCTAGA ACGTGCATCC TTGTAATACC TGATTGATCA 1320
CACAGTGTTA GTGCTGGTCA GAGAGACCTC ATCCTGCTCT TTTGTGATGA ACATATTCAT 1380
GAAATGTGGA AGTCAGTACG ATCAAGTTGT TGACTGTGAT TAGATCACAT CTTAAATTCA 1440
TTTCTAGACT CAAAACCTGG AGATGCAGCT ACTGGAATGG TGTTTTGTCA GACTTCCAAA 1500
TCCTGGAAGG ACACAGTGAT GAATGTACTA TATCTGAACA TAGAAACTCG GGCTTGAGTG 1560
AGAAGAGCTT GCACAGCCAA CGAGACACAT TGCCTTCTGG AGCTGGGAGA CAAAGGAGGA 1620
ATTTACTTTC TTCACCAAGT GCAATAGATT ACTGATGTGA TATTCTGTTG CTTTACAGTT 1680
ACAGTTGATG TTTGGGGATC GATGTGCTCA GCCAAATTTC CTGTTTGAAA TATCATGTTA 1740
AATTAGAATG AATTTATCTT TACCAAAAAC CATGTTGCGT TCAAAGAGGT GAACATTAAA 1800
ATATAGAGAC AGGACAGAAT GTGTTCTTTT CTCCTCTACC AGTCCTATTT TTCAATGGGA 1860
AGACTCAGGA GTCTGCCACT TGTCAAAGAA GGTGCTGATC CTAAGAATTT TTCATTCTCA 1920
GAATTCGGTG TGCTGCCAAC TTGATGTTCC ACCTGCCACA AACCACCAGG ACTGAAAGAA 1980
GAAAACAGTA CAGAAGGCAA AGTTTACAGA TGTTTTTAAT TCTAGTATTT TATCTGGAAC 2040
AACTTGTAGC AGCTATATAT TTCCCCTTGG TCCCAAGCCT GATACTTTAG CCATCATAAC 2100
TCACTAACAG GGAGAAGTAG CTAGTAGCAA TGTGCCTTGA TTGATTAGAT AAAGATTTCT 2160
AGTAGGCAGC AAAAGACCAA ATCTCAGTTG TTTGCTTCTT GCCATCACTG GTCCAGGTCT 2220
TCAGTTTCCG AATCTCTTTC CCTTCCCCTG TGGTCTATTG TCGCTATGTG ACTTGCGCTT 2280
AATCCAATAT TTTGCCTTTT TTCTATATCA AAAAACCTTT ACAGTTAGCA GGGATGTTCC 2340
TTACCGAGGA TTTTTAACCC CCAATCTCTC ATAATCGCTA GTGTTTAAAA GGCTAAGAAT 2400
AGTGGGGCCC AACCGATGTG GTAGGTGATA AAGAGGCATC TTTTCTAGAG ACACATTGGA 2460
CCAGATGAGG ATCCGAAACG GCAGCCTTTA CGTTCATCAC CTGCTAGAAC CTCTCGTAGT 2520
CCATCACCAT TTCTTGGCAT TGGAATTCTA CTGGAAAAAA ATACAAAAAG CAAAACAAAA 2580
CCCTCAGCAC TGTTACAAGA GGCCATTTAA GTATCTTGTG CTTCTTCACT TACCCATTAG 2640
CCAGGTTCTC ATTAGGTTTT GCTTGGGCCT CCCTGGCACT GAACCTTAGG CTTTGTATGA 2700
CAGTGAAGCA GCACTGTGAG TGGTTCAAGC ACACTGGAAT ATAAAACAGT CATGGCCTGA 2760
GATGCAGGTG ATGCCATTAC AGAACCAAAT CGTGGCACGT ATTGCTGTGT CTCCTCTCAG 2820
AGTGACAGTC ATAAATACTG TCAAACAATA AAGGGAGAAT GGTGCTGTTT AAAGTCACAT 2880
CCCTGTAAAT TGCAGAATTC AAAAGTGATT ATCTCTTTGA TCTACTTGCC TCATTTCCCT 2940
ATCTTCTCCC CCACGGTATC CTAAACTTTA GACTTCCCAC TGTTCTGAAA GGAGACATTG 3000
CTCTATGTCT GCCTTCGACC ACAGCAAGCC ATCATCCTCC ATTGCTCCCG GGGACTCAAG 3060
AGGAATCTGT TTCTCTGCTG TCAACTTCCC ATCTGGCTCA GCATAGGGTC ACTTTGCCAT 3120
TATGCAAATG GAGATAAAAG CAATTCTGGC TGTCCAGGAG CTAATCTGAC CGTTCTATTG 3180
TGTGGATGAC CACATAAGAA GGCAATTTTA GTGTATTAAT CATAGATTAT TATAAACTAT 3240
AAACTTAAGG GCAAGGAGTT TATTACAATG TATCTTTATT AAAACAAAAG GGTGTATAGT 3300
GTTCACAAAC TGTGAAAATA GTGTAAGAAC TGTACATTGT GAGCTCTGGT TATTTTTCTC 3360
TTGTACCATA GAAAAATGTA TAAAAATTAT CAAAAAGCTA ATGTGCAGGG ATATTGCCTT 3420
ATTTGTCTGT AAAAAATGGA GCTCAGTAAC ATAACTGCTT CTTGGAGCTT TGGAATATTT 3480
TATCCTGTAT TCTTGTTT 3498

(2) INFORMATION FOR SEQ ID NO:6:
  (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: ~~344~~ 343 amino acids
    (B) TYPE: peptide
    (D) TOPOLOGY: linear
  (ii) MOLECULE TYPE: peptide
  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
                  5                  10                  15
Met Gln Gly Lys Arg Lys Ala Leu Lys Leu Asn Phe Ala Asn Pro Pro
             20                  25                  30
Phe Lys Ser Thr Ala Arg Phe Thr Leu Asn Pro Asn Pro Thr Gly Val
         35                  40                  45
Gln Asn Pro His Ile Glu Arg Leu Arg Thr His Ser Ile Glu Ser Ser
     50                  55                  60
Gly Lys Leu Lys Ile Ser Pro Glu Gln His Trp Asp Phe Thr Ala Glu
 65                  70                  75                  80
Asp Leu Lys Asp Leu Gly Glu Ile Gly Arg Gly Ala Tyr Gly Ser Val
                 85                  90                  95
Asn Lys Met Val His Lys Pro Ser Gly Gln Ile Met Ala Val Lys Arg
            100                 105                 110
Ile Arg Ser Thr Val Asp Glu Lys Glu Gln Lys Gln Leu Leu Met Asp
        115                 120                 125
Leu Asp Val Val Met Arg Ser Ser Asp Cys Pro Tyr Ile Val Gln Phe
    130                 135                 140
Tyr Gly Ala Leu Phe Arg Glu Gly Asp Cys Trp Ile Cys Met Glu Leu
145                 150                 155                 160
Met Ser Thr Ser Phe Asp Lys Phe Tyr Lys Tyr Val Tyr Ser Val Leu
                165                 170                 175
Asp Asp Val Ile Pro Glu Glu Ile Leu Gly Lys Ile Thr Leu Ala Thr
            180                 185                 190
Val Lys Ala Leu Asn His Leu Lys Glu Asn Leu Lys Ile Ile His Arg
        195                 200                 205
Asp Ile Lys Pro Ser Asn Ile Leu Leu Asp Arg Ser Gly Asn Ile Lys
    210                 215                 220
Leu Cys Asp Phe Gly Ile Ser Gly Gln Leu Val Asp Ser Ile Ala Lys
225                 230                 235                 240
Thr Arg Asp Ala Gly Cys Arg Pro Tyr Met Ala Pro Glu Arg Ile Asp
                245                 250                 255
Pro Ser Ala Ser Arg Gln Gly Tyr Asp Val Arg Ser Asp Val Trp Ser
            260                 265                 270
Leu Gly Ile Thr Leu Tyr Glu Leu Ala Thr Gly Arg Phe Pro Tyr Pro
        275                 280                 285
Lys Trp Asn Ser Val Phe Asp Gln Leu Thr Gln Val Val Lys Gly Asp
    290                 295                 300
Pro Pro Gln Leu Ser Asn Ser Glu Glu Arg Glu Phe Ser Pro Ser Phe
305                 310                 315                 320
Ile Asn Phe Val Asn Leu Cys Leu Thr Lys Asp Glu Ser Lys Arg Pro
                325                 330                 335
Lys Tyr Lys Glu Leu Leu Lys His Pro Phe Ile Leu Met Tyr Glu Glu
            340                 345                 350
Arg Ala Val Glu Val Ala Cys Tyr Val Cys Lys Ile Leu Asp Gln Met
        355                 360                 363
Pro Ala Thr Pro Ser Ser Pro Met Tyr Val Asp
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3554 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CAACAATGGC GGCTCCGAGC CCGAGCGGTG GCGGCGGCAG CGGCACCCCC GGCCCCGTAG 60
GGTCCCCGGC GCCAGGCCAC CCGGCCGTCA GCAGCATGCA GGGTAAACGC AAAGCACTGA 120
AGTTGAATTT TGCAAATCCA CCTTTCAAAT CTACAGCAAG GTTTACTCTG AATCCCAATC 180
CTACAGGAGT TCAAAACCCA CACATAGAGA GACTGAGAAC ACACAGCATT GAGTCATCAG 240
GAAAACTGAA GATCTCCCCT GAACAACACT GGGATTTCAC TGCAGAGGAC TTGAAAGACC 300
TTGGAGAAAT TGGACGAGGA GCTTATGGTT CTGTCAACAA AATGGTCCAC AAACCAAGTG 360
GGCAAATAAT GGCAGTTAAA AGAATTCGGT CAACAGTGGA TGAAAAAGAA CAAAAACAAC 420
TTCTTATGGA TTTGGATGTA GTAATGCGGA GTAGTGATTG CCCATACATT GTTCAGTTTT 480
ATGGTGCACT CTTCAGAGAG GGTGACTGTT GGATCTGTAT GGAACTCATG TCTACCTCGT 540
TTGATAAGTT TTACAAATAT GTATATAGTG TATTAGATGA TGTTATTCCA GAAGAAATTT 600
TAGGCAAAAT CACTTTAGCA ACTGTGAAAG CACTAAACCA CTTAAAAGAA AACTTGAAAA 660
TTATTCACAG AGATATCAAA CCTTCCAATA TTCTTCTGGA CAGAAGTGGA AATATTAAGC 720
TCTGTGACTT CGGCATCAGT GGACAGCTTG TGGACTCTAT TGCCAAGACA AGAGATGCTG 780
GCTGTAGGCC ATACATGGCA CCTGAAAGAA TAGACCCAAG CGCATCACGA CAAGGATATG 840
ATGTCCGCTC TGATGTCTGG AGTTTGGGGA TCACATTGTA TGAGTTGGCC ACAGGCCGAT 900
TTCCTTATCC AAAGTGGAAT AGTGTATTTG ATCAACTAAC ACAAGTCGTG AAAGGAGATC 960
CTCCGCAGCT GAGTAATTCT GAGGAAAGGG AATTCTCCCC GAGTTTCATC AACTTTGTCA 1020
ACTTGTGCCT TACGAAGGAT GAATCCAAAA GGCCAAAGTA TAAAGAGCTT CTGAAACATC 1080
CCTTTATTTT GATGTATGAA GAACGTGCCG TTGAGGTCGC ATGCTATGTT TGTAAAATCC 1140
TGGATCAAAT GCCAGCTACT CCCAGCTCTC CCATGTATGT CGATTGATAT CGYTGCTACA 1200
TCAGACTCTA GAAAAAAGGG CTGAGAGGAA GCAAGACGTA AAGAATTTTC ATCCCGTATC 1260
ACAGTGTTTT TATTGCTCGC CCAGACACCA TGTGCAATAA GATTGGTGTT CGTTTCCATC 1320
ATGTCTGTAT ACTCCTGTCA CCTAGAACGT GCATCCTTGT AATACCTGAT TGATCACACA 1380
GTGTTAGTGC TGGTCAGAGA GACCTCATCC TGCTCTTTTG TGATGAACAT ATTCATGAAA 1440
TGTGGAAGTC AGTACGATCA AGTTGTTGAC TGTGATTAGA TCACATCTTA AATTCATTTC 1500
TAGACTCAAA ACCTGGAGAT GCAGCTACTG GAATGGTGTT TTGTCAGACT TCCAAATCCT 1560
GGAAGGACAC AGTGATGAAT GTACTATATC TGAACATAGA AACTCGGGCT TGAGTGAGAA 1620
GAGCTTGCAC AGCCAACGAG ACACATTGCC TTCTGGAGCT GGGAGACAAA GGAGGAATTT 1680
ACTTTCTTCA CCAAGTGCAA TAGATTACTG ATGTGATATT CTGTTGCTTT ACAGTTACAG 1740
TTGATGTTTG GGGATCGATG TGCTCAGCCA AATTTCCTGT TTGAAATATC ATGTTAAATT 1800
AGAATGAATT TATCTTTACC AAAAACCATG TTGCGTTCAA AGAGGTGAAC ATTAAAATAT 1860
AGAGACAGGA CAGAATGTGT TCTTTTCTCC TCTACCAGTC CTATTTTTCA ATGGGAAGAC 1920
TCAGGAGTCT GCCACTTGTC AAAGAAGGTG CTGATCCTAA GAATTTTTCA TTCTCAGAAT 1980
TCGGTGTGCT GCCAACTTGA TGTTCCACCT GCCACAAACC ACCAGGACTG AAAGAAGAAA 2040
ACAGTACAGA AGGCAAAGTT TACAGATGTT TTTAATTCTA GTATTTTATC TGGAACAACT 2100
TGTAGCAGCT ATATATTTCC CCTTGGTCCC AAGCCTGATA CTTTAGCCAT CATAACTCAC 2160
TAACAGGGAG AAGTAGCTAG TAGCAATGTG CCTTGATTGA TTAGATAAAG ATTTCTAGTA 2220
GGCAGCAAAA GACCAAATCT CAGTTGTTTG CTTCTTGCCA TCACTGGTCC AGGTCTTCAG 2280
TTTCCGAATC TCTTTCCCTT CCCCTGTGGT CTATTGTCGC TATGTGACTT GCGCTTAATC 2340
CAATATTTTG CCTTTTTTCT ATATCAAAAA ACCTTTACAG TTAGCAGGGA TGTTCCTTAC 2400
CGAGGATTTT TAACCCCCAA TCTCTCATAA TCGCTAGTGT TTAAAAGGCT AAGAATAGTG 2460
GGGCCCAACC GATGTGGTAG GTGATAAAGA GGCATCTTTT CTAGAGACAC ATTGGACCAG 2520
ATGAGGATCC GAAACGGCAG CCTTTACGTT CATCACCTGC TAGAACCTCT CGTAGTCCAT 2580

```
CACCATTTCT TGGCATTGGA ATTCTACTGG AAAAAAATAC AAAAAGCAAA ACAAAACCCT 2640
CAGCACTGTT ACAAGAGGCC ATTTAAGTAT CTTGTGCTTC TTCACTTACC CATTAGCCAG 2700
GTTCTCATTA GGTTTTGCTT GGGCCTCCCT GGCACTGAAC CTTAGGCTTT GTATGACAGT 2760
GAAGCAGCAC TGTGAGTGGT TCAAGCACAC TGGAATATAA AACAGTCATG GCCTGAGATG 2820
CAGGTGATGC CATTACAGAA CCAAATCGTG GCACGTATTG CTGTGTCTCC TCTCAGAGTG 2880
ACAGTCATAA ATACTGTCAA ACAATAAAGG GAGAATGGTG CTGTTTAAAG TCACATCCCT 2940
GTAAATTGCA GAATTCAAAA GTGATTATCT CTTTGATCTA CTTGCCTCAT TTCCCTATCT 3000
TCTCCCCCAC GGTATCCTAA ACTTTAGACT TCCCACTGTT CTGAAAGGAG ACATTGCTCT 3060
ATGTCTGCCT TCGACCACAG CAAGCCATCA TCCTCCATTG CTCCCGGGGA CTCAAGAGGA 3120
ATCTGTTTCT CTGCTGTCAA CTTCCCATCT GGCTCAGCAT AGGGTCACTT TGCCATTATG 3180
CAAATGGAGA TAAAAGCAAT TCTGGCTGTC CAGGAGCTAA TCTGACCGTT CTATTGTGTG 3240
GATGACCACA TAAGAAGGCA ATTTTAGTGT ATTAATCATA GATTATTATA AACTATAAAC 3300
TTAAGGGCAA GGAGTTTATT ACAATGTATC TTTATTAAAA CAAAAGGGTG TATAGTGTTC 3360
ACAAACTGTG AAAATAGTGT AAGAACTGTA CATTGTGAGC TCTGGTTATT TTTCTCTTGT 3420
ACCATAGAAA AATGTATAAA AATTATCAAA AAGCTAATGT GCAGGGATAT TGCCTTATTT 3480
GTCTGTAAAA AATGGAGCTC AGTAACATAA CTGCTTCTTG GAGCTTTGGA ATATTTTATC 3540
CTGTATTCTT GTTT                                                   3554
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: ~~404~~ 412 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
                  5                  10                  15
Met Ala Ala Pro Ser Pro Ser Gly Gly Gly Gly Ser Gly Thr Pro Gly
            20                  25                  30
Pro Val Gly Ser Pro Ala Pro Gly His Pro Ala Val Ser Ser Met Gln
        35                  40                  45
Gly Lys Arg Lys Ala Leu Lys Leu Asn Phe Ala Asn Pro Pro Phe Lys
    50                  55                  60
Ser Thr Ala Arg Phe Thr Leu Asn Pro Asn Pro Thr Gly Val Gln Asn
65                  70                  75                  80
Pro His Ile Glu Arg Leu Arg Thr His Ser Ile Glu Ser Ser Gly Lys
                85                  90                  95
Leu Lys Ile Ser Pro Glu Gln His Trp Asp Phe Thr Ala Glu Asp Leu
            100                 105                 110
Lys Asp Leu Gly Glu Ile Gly Arg Gly Ala Tyr Gly Ser Val Asn Lys
        115                 120                 125
Met Val His Lys Pro Ser Gly Gln Ile Met Ala Val Lys Arg Ile Arg
    130                 135                 140
Ser Thr Val Asp Glu Lys Glu Gln Lys Gln Leu Leu Met Asp Leu Asp
145                 150                 155                 160
Val Val Met Arg Ser Ser Asp Cys Pro Tyr Ile Val Gln Phe Tyr Gly
                165                 170                 175
Ala Leu Phe Arg Glu Gly Asp Cys Trp Ile Cys Met Glu Leu Met Ser
            180                 185                 190
Thr Ser Phe Asp Lys Phe Tyr Lys Tyr Val Tyr Ser Val Leu Asp Asp
```

```
        195                 200                 205
Val Ile Pro Glu Glu Ile Leu Gly Lys Ile Thr Leu Ala Thr Val Lys
    210                 215                 220
Ala Leu Met His Leu Lys Glu Asn Leu Lys Ile Ile His Arg Asp Ile
225                 230                 235                 240
Lys Pro Ser Asn Ile Leu Leu Asp Arg Ser Gly Met Ile Lys Leu Cys
                245                 250                 255
Asp Phe Gly Ile Ser Gly Gln Leu Val Asp Ser Ile Ala Lys Thr Arg
            260                 265                 270
Asp Ala Gly Cys Arg Pro Tyr Met Ala Pro Glu Arg Ile Asp Phe Ser
        275                 280                 285
Ala Ser Arg Gln Gly Tyr Asp Val Arg Ser Asp Val Trp Ser Leu Gly
    290                 295                 300
Ile Thr Leu Tyr Glu Leu Ala Thr Gly Arg Phe Pro Tyr Pro Lys Trp
305                 310                 315                 320
Asn Ser Val Phe Asp Gln Leu Thr Gln Val Val Lys Gly Asp Pro Pro
                325                 330                 335
Gln Leu Ser Asn Ser Glu Glu Arg Glu Phe Ser Pro Ser Phe Ile Asn
            340                 345                 350
Phe Val Asn Leu Cys Leu Thr Lys Asp Glu Ser Lys Arg Pro Lys Tyr
        355                 360                 365
Lys Glu Leu Leu Lys His Pro Phe Ile Leu Met Tyr Glu Glu Arg Ala
    370                 375                 380
Val Glu Val Ala Cys Tyr Val Cys Lys Ile Leu Asp Gln Met Pro Ala
385                 390                 393
Thr Pro Ser Ser Pro Met Tyr Val Asp
```

(2) INFORMATION FOR SEQ ID NO:9

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3636 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CTCCCAACAA TGGCGGCTCC GAGCCCGAGC GGCGGCGGCG GCTCCGGGGG CGGCAGCGGC   60
AGCGGCACCC CCGGCCCCGT AGGGTCCCCG GCGCCAGGCC ACCCGGCCGT CAGCAGCATG  120
CAGGGTAAAC GCAAAGCACT GAAGTTGAAT TTTGCAAATC CACCTTTCAA ATCTACAGCA  180
AGGTTTACTC TGAATCCCAA TCCTACAGGA GTTCAAAACC CACACATAGA GAGACTGAGA  240
ACACACAGCA TTGAGTCATC AGGAAAACTG AAGATCTCCC CTGAACAACA CTGGGATTTC  300
ACTGCAGAGG ACTTGAAAGA CCTTGGAGAA ATTGGACGAG GAGCTTATGG TTCTGTCAAC  360
AAAATGGTCC ACAAACCAAG TGGGCAAATA ATGGCAGTTA AAAGAATTCG GTCAACAGTG  420
GATGAAAAAG AACAAAAACA ACTTCTTATG GATTTGGATG TAGTAATGCG GAGTAGTGAT  480
TGCCCATACA TTGTTCAGTT TTATGGTGCA CTCTTCAGAG AGGGTGACTG TTGGATCTGT  540
ATGGAACTCA TGTCTACCTC GTTTGATAAG TTTTACAAAT ATGTATATAG TGTATTAGAT  600
GATGTTATTC CAGAAGAAAT TTTAGGCAAA ATCACTTTAG CAACTGTGAA AGCACTAAAC  660
CACTTAAAAG AAAACTTGAA AATTATTCAC AGAGATATCA AACCTTCCAA TATTCTTCTG  720
GACAGAAGTG GAAATATTAA GCTCTGTGAC TTCGGCATCA GTGGACAGCT TGTGGACTCT  780
ATTGCCAAGA CAAGAGATGC TGGCTGTAGG CCATACATGG CACCTGAAAG AATAGACCCA  840
```

AGCGCATCAC GACAAGGATA TGATGTCCGC TCTGATGTCT GGAGTTTGGG GATCACATTG 900
TATGAGTTGG CCACAGGCCG ATTTCCTTAT CCAAAGTGGA ATAGTGTATT TGATCAACTA 960
ACACAAGTCG TGAAAGGAGA TCCTCCGCAG CTGAGTAATT CTGAGGAAAG GGAATTCTCC 1020
CCGAGTTTCA TCAACTTTGT CAACTTGTGC CTTACGAAGG ATGAATCCAA AAGGCCAAAG 1080
TATAAAGAGC TTCTGAAACA TCCCTTTATT TTGATGTATG AAGAACGTGC CGTTGAGGTC 1140
GCATGCTATG TTTGTAAAAT CCTGGATCAA ATGCCAGCTA CTCCCAGCTC TCCCATGTAT 1200
GTCGATTGAT ATCGCTGCTA CATCAGACTC TAGAAAAAAG GGCTGAGAGG AAGCAAGACG 1260
TAAAGAATTT TCATCCCGTA TCACAGTGTT TTTATTGCTC GCCCAGACAC CATGTGCAAT 1320
AAGATTGGTG TTCGTTTCCA TCATGTCTGT ATACTCCTGT CACCTAGAAC GTGCATCCTT 1380
GTAATACCTG ATTGATCACA CAGTGTTAGT GCTGGTCAGA GAGACCTCAT CCTGCTCTTT 1440
TGTGATGAAC ATATTCATGA AATGTGGAAG TCAGTACGAT CAAGTTGTTG ACTGTGATTA 1500
GATCACATCT TAAATTCATT TCTAGACTCA AAACCTGGAG ATGCAGCTAC TGGAATGGTG 1560
TTTTGTCAGA CTTCCAAATC CTGGAAGGAC ACAGTGATGA ATGTACTATA TCTGAACATA 1620
GAAACTCGGG CTTGAGTGAG AAGAGCTTGC ACAGCCAACG AGACACATTG CCTTCTGGAG 1680
CTGGGAGACA AAGGAGGAAT TTACTTTCTT CACCAAGTGC AATAGATTAC TGATGTGATA 1740
TTCTGTTGCT TTACAGTTAC AGTTGATGTT TGGGGATCGA TGTGCTCAGC CAAATTTCCT 1800
GTTTGAAATA TCATGTTAAA TTAGAATGAA TTTATCTTTA CCAAAAACCA TGTTGCGTTC 1860
AAAGAGGTGA ACATTAAAAT ATAGAGACAG GACAGAATGT GTTCTTTTCT CCTCTACCAG 1920
TCCTATTTTT CAATGGGAAG ACTCAGGAGT CTGCCACTTG TCAAAGAAGG TGCTGATCCT 1980
AAGAATTTTT CATTCTCAGA ATTCGGTGTG CTGCCAACTT GATGTTCCAC CTGCCACAAA 2040
CCACCAGGAC TGAAAGAAGA AAACAGTACA GAAGGCAAAG TTTACAGATG TTTTTAATTC 2100
TAGTATTTTA TCTGGAACAA CTTGTAGCAG CTATATATTT CCCCTTGGTC CCAAGCCTGA 2160
TACTTTAGCC ATCATAACTC ACTAACAGGG AGAAGTAGCT AGTAGCAATG TGCCTTGATT 2220
GATTAGATAA AGATTTCTAG TAGGCAGCAA AAGACCAAAT CTCAGTTGTT TGCTTCTTGC 2340
CATCACTGGT CCAGGTCTTC AGTTTCCGAA TCTCTTTCCC TTCCCCTGTG GTCTATTGTC 2400
GCTATGTGAC TTGCGCTTAA TCCAATATTT TGCCTTTTTT CTATATCAAA AAACCTTTAC 2460
AGTTAGCAGG GATGTTCCTT ACCGAGGATT TTTAACCCCC AATCTCTCAT AATCGCTAGT 2520
GTTTAAAAGG CTAAGAATAG TGGGGCCCAA CCGATGTGGT AGGTGATAAA GAGGCATCTT 2580
TTCTAGAGAC ACATTGGACC AGATGAGGAT CCGAAACGGC AGCCTTTACG TTCATCACCT 2640
GCTAGAACCT CTCGTAGTCC ATCACCATTT CTTGGCATTG GAATTCTACT GGAAAAAAAT 2700
ACAAAAAGCA AAACAAAACC CTCAGCACTG TTACAAGAGG CCATTTAAGT ATCTTGTGCT 2760
TCTTCACTTA CCCATTAGCC AGGTTCTCAT TAGGTTTTGC TTGGGCCTCC CTGGCACTGA 2820
ACCTTAGGCT TTGTATGACA GTGAAGCAGC ACTGTGAGTG GTTCAAGCAC ACTGGAATAT 2880
AAAACAGTCA TGGCCTGAGA TGCAGGTGAT GCCATTACAG AACCAAATCG TGGCACGTAT 2940
TGCTGTGTCT CCTCTCAGAG TGACAGTCAT AAATACTGTC AAACAATAAA GGGAGAATGG 3000
TGCTGTTTAA AGTCACATCC CTGTAAATTG CAGAATTCAA AAGTGATTAT CTCTTTGATC 3060
TACTTGCCTC ATTTCCCTAT CTTCTCCCCC ACGGTATCCT AAACTTTAGA CTTCCCACTG 3120
TTCTGAAAGG AGACATTGCT CTATGTCTGC CTTCGACCAC AGCAAGCCAT CATCCTCCAT 3180
TGCTCCCGGG GACTCAAGAG GAATCTGTTT CTCTGCTGTC AACTTCCCAT CTGGCTCAGC 3240
ATAGGGTCAC TTTGCCATTA TGCAAATGGA GATAAAAGCA ATTCTGGCTG TCCAGGAGCT 3300
AATCTGACCG TTCTATTGTG TGGATGACCA CATAAGAAGG CAATTTTAGT GTATTAATCA 3360
TAGATTATTA TAAACTATAA ACTTAAGGGC AAGGAGTTTA TTACAATGTA TCTTTATTAA 3420
AACAAAAGGG TGTATAGTGT TCACAAACTG TGAAAATAGT GTAAGAACTG TACATTGTGA 3480
GCTCTGGTTA TTTTTCTCTT GTACCATAGA AAAATGTATA AAAATTATCA AAAAGCTAAT 3540
GTGCAGGGAT ATTGCCTTAT TTGTCTGTAA AAAATGGAGC TCAGTAACAT AACTGCTTCT 3600
TGGAGCTTTG GAATATTTTA TCCTGTATTC TTGTTT 3636

(2) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 399 amino acids
(B) TYPE: peptide
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
                  5                  10                  15
Met Ala Ala Pro Ser Pro Ser Gly Gly Gly Gly Ser Gly Gly Gly Ser
            20                  25                  30
Gly Ser Gly Thr Pro Gly Pro Val Gly Ser Pro Ala Pro Gly His Pro
        35                  40                  45
Ala Val Ser Ser Met Gln Gly Lys Arg Lys Ala Leu Lys Leu Asn Phe
    50                  55                  60
Ala Asn Pro Pro Phe Lys Ser Thr Ala Arg Phe Thr Leu Asn Pro Asn
65                  70                  75                  80
Pro Thr Gly Val Gln Asn Pro His Ile Glu Arg Leu Arg Thr His Ser
                85                  90                  95
Ile Glu Ser Ser Gly Lys Leu Lys Ile Ser Pro Glu Gln His Trp Asp
            100                 105                 110
Phe Thr Ala Glu Asp Leu Lys Asp Leu Gly Glu Ile Gly Arg Gly Ala
        115                 120                 125
Tyr Gly Ser Val Asn Lys Met Val His Lys Pro Ser Gly Gln Ile Met
    130                 135                 140
Ala Val Lys Arg Ile Arg Ser Thr Val Asp Glu Lys Glu Gln Lys Gln
145                 150                 155                 160
Leu Leu Met Asp Leu Asp Val Val Met Arg Ser Ser Asp Cys Pro Tyr
                165                 170                 175
Ile Val Gln Phe Tyr Gly Ala Leu Phe Arg Glu Gly Asp Cys Trp Ile
            180                 185                 190
Cys Met Glu Leu Met Ser Thr Ser Phe Asp Lys Phe Tyr Lys Tyr Val
        195                 200                 205
Tyr Ser Val Leu Asp Asp Val Ile Pro Glu Glu Ile Leu Gly Lys Ile
    210                 215                 220
Thr Leu Ala Thr Val Lys Ala Leu Asn His Leu Lys Glu Asn Leu Lys
225                 230                 235                 240
Ile Ile His Arg Asp Ile Lys Pro Ser Asn Ile Leu Leu Asp Arg Ser
                245                 250                 255
Gly Asn Ile Lys Leu Cys Asp Phe Gly Ile Ser Gly Gln Leu Val Asp
            260                 265                 270
Ser Ile Ala Lys Thr Arg Asp Ala Gly Cys Arg Pro Tyr Met Ala Pro
        275                 280                 285
Glu Arg Ile Asp Pro Ser Ala Ser Arg Gln Gly Tyr Asp Val Arg Ser
    290                 295                 300
Asp Val Trp Ser Leu Gly Ile Thr Leu Tyr Glu Leu Ala Thr Gly Arg
305                 310                 315                 320
Phe Pro Tyr Pro Lys Trp Asn Ser Val Phe Asp Gln Leu Thr Gln Val
                325                 330                 335
Val Lys Gly Asp Pro Pro Gln Leu Ser Asn Ser Glu Glu Arg Glu Phe
            340                 345                 350
Ser Pro Ser Phe Ile Asn Phe Val Asn Leu Cys Leu Thr Lys Asp Glu
```

```
        355                 360                 365
Ser Lys Arg Pro Lys Tyr Lys Glu Leu Leu Lys His Pro Phe Ile Leu
    370                 375                 380
Met Tyr Glu Glu Arg Ala Val Glu Val Ala Cys Tyr Val Cys Lys Ile
385                 390                 395                 399
Leu Asp Gln Met Pro Ala Thr Pro Ser Ser Pro Met Tyr Val Asp
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 393 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
                5                  10                  15
Met Pro Lys Lys Lys Pro Thr Pro Ile Gln Leu Asn Pro Ala Pro Asp
            20                  25                  30
Gly Ser Ala Val Asn Gly Thr Ser Ser Ala Glu Thr Asn Leu Glu Ala
        35                  40                  45
Leu Gln Lys Lys Leu Glu Glu Leu Glu Leu Asp Glu Gln Gln Arg Lys
    50                  55                  60
Arg Leu Glu Ala Phe Leu Thr Gln Lys Gln Lys Val Gly Glu Leu Lys
65                  70                  75                  80
Asp Asp Asp Phe Glu Lys Ile Ser Glu Leu Gly Ala Gly Asn Gly Gly
                85                  90                  95
Val Val Phe Lys Val Ser His Lys Pro Ser Gly Leu Val Met Ala Arg
            100                 105                 110
Lys Leu Ile His Leu Glu Ile Lys Pro Ala Ile Arg Asn Gln Ile Ile
        115                 120                 125
Arg Glu Leu Gln Val Leu His Glu Cys Asn Ser Pro Tyr Ile Val Gly
    130                 135                 140
Phe Tyr Gly Ala Phe Tyr Ser Asp Gly Glu Ile Ser Ile Cys Met Glu
145                 150                 155                 160
His Met Asp Gly Gly Ser Leu Asp Gln Val Leu Lys Lys Ala Gly Arg
                165                 170                 175
Ile Pro Glu Gln Ile Leu Gly Lys Val Ser Ile Ala Val Ile Lys Gly
            180                 185                 190
Leu Thr Tyr Leu Arg Glu Lys His Lys Ile Met His Arg Asp Val Lys
        195                 200                 205
Pro Ser Asn Ile Leu Val Asn Ser Arg Gly Glu Ile Lys Leu Cys Asp
    210                 215                 220
Phe Gly Val Ser Gly Gln Leu Ile Asp Ser Met Ala Asn Ser Phe Val
225                 230                 235                 240
Gly Thr Arg Ser Tyr Met Ser Pro Glu Arg Leu Gln Gly Thr His Tyr
                245                 250                 255
Ser Val Gln Ser Asp Ile Trp Ser Met Gly Leu Ser Leu Val Glu Met
            260                 265                 270
Ala Val Gly Arg Tyr Pro Ile Pro Pro Pro Asp Ala Lys Glu Leu Glu
```

```
        275                 280                 285
Leu Met Phe Gly Cys Gln Val Glu Gly Asp Ala Ala Glu Thr Pro Pro
    290                 295                 300
Arg Pro Arg Thr Pro Gly Arg Pro Leu Ser Ser Tyr Gly Met Asp Ser
305                 310                 315                 320
Arg Pro Pro Met Ala Ile Phe Glu Leu Leu Asp Tyr Ile Val Asn Glu
                325                 330                 335
Pro Pro Pro Lys Leu Pro Ser Gly Val Phe Ser Leu Glu Phe Gln Asp
            340                 345                 350
Phe Val Asn Lys Cys Leu Ile Lys Asn Pro Ala Glu Arg Ala Asp Leu
        355                 360                 365
Lys Gln Leu Met Val His Ala Phe Ile Lys Arg Ser Asp Ala Glu Glu
    370                 375                 380
Val Asp Phe Ala Gly Trp Leu Cys Ser Thr Ile Gly Leu Asn Gln Pro
385                 390                 393
Ser Thr Pro Thr His Ala Ala Gly Val
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 400 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
                5                   10                  15
Met Leu Ala Arg Arg Lys Pro Val Leu Pro Ala Leu Thr Ile Asn Pro
            20                  25                  30
Thr Ile Ala Glu Gly Pro Ser Pro Thr Ser Glu Gly Ala Ser Glu Ala
        35                  40                  45
Asn Leu Val Asp Leu Gln Lys Lys Leu Glu Glu Leu Glu Leu Asp Glu
    50                  55                  60
Gln Gln Lys Lys Arg Leu Glu Ala Phe Leu Thr Gln Lys Ala Lys Val
65                  70                  75                  80
Ser Glu Leu Lys Asp Asp Asp Phe Glu Arg Ile Ser Glu Leu Gly Ala
                85                  90                  95
Gly Asn Gly Gly Val Val Thr Lys Val Gln His Arg Pro Ser Gly Leu
            100                 105                 110
Ile Met Ala Arg Lys Leu Ile His Leu Glu Ile Lys Pro Ala Ile Arg
        115                 120                 125
Asn Gln Ile Ile Arg Glu Leu Gln Val Leu His Glu Cys Asn Ser Pro
    130                 135                 140
Tyr Ile Val Gly Phe Tyr Gly Ala Phe Tyr Ser Asp Gly Glu Ile Ser
145                 150                 155                 160
Ile Cys Met Glu His Met Asp Gly Gly Ser Leu Asp Gln Val Leu Lys
                165                 170                 175
Glu Ala Lys Arg Ile Pro Glu Glu Ile Leu Gly Lys Val Ser Ile Ala
            180                 185                 190
Val Leu Arg Gly Leu Ala Tyr Leu Arg Glu Lys His Gln Ile Met His
```

```
        195                 200                 205
Arg Asp Val Lys Pro Ser Asn Ile Leu Val Asn Ser Arg Gly Glu Ile
    210                 215                 220
Lys Leu Cys Asp Phe Gly Val Ser Gly Gln Leu Ile Asp Ser Met Ala
225                 230                 235                 240
Asn Ser Phe Val Gly Thr Arg Ser Tyr Met Ala Pro Glu Arg Leu Gln
                245                 250                 255
Gly Thr His Tyr Ser Val Gln Ser Asp Ile Trp Ser Met Gly Leu Ser
            260                 265                 270
Leu Val Glu Leu Ala Val Gly Arg Tyr Pro Ile Pro Pro Pro Asp Ala
        275                 280                 285
Lys Glu Leu Glu Ala Ile Phe Gly Arg Pro Val Val Asp Gly Glu Glu
    290                 295                 300
Gly Glu Pro His Ser Ile Ser Pro Arg Pro Arg Pro Pro Gly Arg Pro
305                 310                 315                 320
Val Ser Gly His Gly Met Asp Ser Arg Pro Ala Met Ala Ile Phe Glu
                325                 330                 335
Leu Leu Asp Tyr Ile Val Asn Glu Pro Pro Pro Lys Leu Pro Asn Gly
            340                 345                 350
Val Phe Thr Pro Asp Phe Gln Glu Phe Val Asn Lys Cys Leu Ile Lys
        355                 360                 365
Asn Pro Ala Glu Arg Ala Asp Leu Lys Met Leu Thr Asn His Thr Phe
    370                 375                 380
Ile Lys Arg Ser Glu Val Glu Glu Val Asp Phe Ala Gly Trp Leu Cys
385                 390                 395                 400
Lys Thr Leu Arg Leu Asn Gln Pro Gly Thr Pro Thr Arg Thr Ala Val
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 668 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
                5                   10                  15
Met Glu Asp Lys Phe Ala Asn Leu Ser Leu His Glu Lys Thr Gly Lys
            20                  25                  30
Ser Ser Ile Gln Leu Asn Glu Gln Thr Gly Ser Asp Asn Gly Ser Ala
        35                  40                  45
Val Lys Arg Thr Ser Ser Thr Ser Ser His Tyr Asn Asn Ile Asn Ala
    50                  55                  60
Asp Leu His Ala Arg Val Lys Ala Phe Gln Glu Gln Arg Ala Leu Lys
65                  70                  75                  80
Arg Ser Ala Ser Val Gly Ser Asn Gln Ser Glu Gln Asp Lys Gly Ser
                85                  90                  95
Ser Gln Ser Pro Lys His Ile Gln Gln Ile Val Asn Lys Pro Leu Pro
            100                 105                 110
Pro Leu Pro Val Ala Gly Ser Ser Lys Val Ser Gln Arg Met Ser Ser
```

```
                115                 120                 125
Gln Val Val Gln Ala Ser Ser Lys Ser Thr Leu Lys Asn Val Leu Asp
    130                 135                 140
Asn Gln Glu Thr Gln Asn Ile Thr Asp Val Asn Ile Asn Ile Asp Thr
145                 150                 155                 160
Thr Lys Ile Thr Ala Thr Thr Ile Gly Val Asn Ile Gly Leu Pro Ala
                165                 170                 175
Thr Asp Ile Thr Pro Ser Val Ser Asn Thr Ala Ser Ala Thr His Lys
            180                 185                 190
Ala Gln Leu Leu Asn Pro Asn Arg Arg Ala Pro Arg Arg Pro Leu Ser
        195                 200                 205
Thr Gln His Pro Thr Arg Pro Asn Val Ala Pro His Lys Ala Pro Ala
    210                 215                 220
Ile Ile Asn Thr Pro Lys Gln Ser Leu Ser Ala Arg Arg Gly Leu Lys
225                 230                 235                 240
Leu Pro Pro Gly Gly Met Ser Leu Lys Met Pro Thr Lys Thr Ala Gln
                245                 250                 255
Gln Pro Gln Gln Phe Ala Pro Ser Pro Ser Asn Lys Lys His Ile Glu
            260                 265                 270
Thr Leu Ser Asn Ser Lys Val Val Glu Gly Lys Arg Ser Asn Pro Gly
        275                 280                 285
Ser Leu Ile Asn Gly Val Gln Ser Thr Ser Thr Ser Ser Ser Thr Glu
    290                 295                 300
Gly Pro His Asp Thr Val Gly Thr Thr Pro Arg Thr Gly Asn Ser Asn
305                 310                 315                 320
Asn Ser Ser Asn Ser Gly Ser Ser Gly Gly Gly Gly Leu Phe Ala Asn
                325                 330                 335
Phe Ser Lys Tyr Val Asp Ile Lys Ser Gly Ser Leu Asn Phe Ala Gly
            340                 345                 350
Lys Leu Ser Leu Ser Ser Lys Gly Ile Asp Phe Ser Asn Gly Ser Ser
        355                 360                 365
Ser Arg Ile Thr Leu Asp Glu Leu Glu Phe Leu Asp Glu Leu Gly His
    370                 375                 380
Gly Asn Tyr Gly Asn Val Ser Lys Val Leu His Lys Pro Thr Asn Val
385                 390                 395                 400
Ile Met Ala Thr Lys Glu Val Arg Leu Glu Leu Asp Glu Ala Lys Phe
                405                 410                 415
Arg Gln Ile Leu Met Glu Leu Glu Val Leu His Lys Cys Asn Ser Pro
            420                 425                 430
Tyr Ile Val Asp Phe Tyr Gly Ala Phe Phe Ile Glu Gly Ala Val Tyr
        435                 440                 445
Met Cys Met Glu Tyr Met Asp Gly Gly Ser Leu Asp Lys Ile Tyr Asp
    450                 455                 460
Glu Ser Ser Glu Ile Gly Gly Ile Asp Glu Pro Gln Leu Ala Phe Ile
465                 470                 475                 480
Ala Asn Ala Val Ile His Gly Leu Lys Glu Leu Lys Glu Gln His Asn
                485                 490                 495
Ile Ile His Arg Asp Val Lys Pro Thr Asn Ile Leu Cys Ser Ala Asn
```

```
                500                 505                 510
Gln Gly Thr Val Lys Leu Cys Asp Phe Gly Val Ser Gly Asn Leu Val
        515                 520                 525
Ala Ser Leu Ala Lys Thr Asn Ile Gly Cys Gln Ser Tyr Met Ala Pro
    530                 535                 540
Glu Arg Ile Lys Ser Leu Asn Pro Asp Arg Ala Thr Tyr Thr Val Gln
545                 550                 555                 560
Ser Asp Ile Trp Ser Leu Gly Leu Ser Ile Leu Glu Met Ala Leu Gly
                565                 570                 575
Arg Tyr Pro Tyr Pro Pro Glu Thr Tyr Asp Asn Ile Phe Ser Gln Leu
            580                 585                 590
Ser Ala Ile Val Asp Gly Pro Pro Pro Arg Leu Pro Ser Asp Lys Phe
        595                 600                 605
Ser Ser Asp Ala Gln Asp Phe Val Ser Leu Cys Leu Gln Lys Ile Pro
    610                 615                 620
Glu Arg Arg Pro Thr Tyr Ala Ala Leu Thr Glu His Pro Trp Leu Val
625                 630                 635                 640
Lys Tyr Arg Asn Gln Asp Val His Met Ser Glu Tyr Ile Thr Glu Arg
                645                 650                 655
Leu Glu Arg Arg Asn Lys Ile Leu Arg Glu Arg Gly Glu Asn Gly Leu
            660                 605                 608
Ser Lys Asn Val Pro Ala Leu His Met Gly Gly Leu
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: ~~2~~13 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTYTAYGGNG CNTTYTTYAT HGA ~~2~~13

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: ~~3636~~ base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATBCTYTCNG GNGCCATKTA 20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: peptide

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Asp Tyr Lys Asp Asp Asp Asp Lys
5 8